Webを支える情報通信技術（7）
――大事なコンテンツを守るには――

2007年6月4日
情報基盤センター
佐藤周行・中山雅哉・西村健

１．会員限定のコンテンツ

コンテンツを作成したら、Webサーバを使って公開することができる。Webサーバの挙動については、何回かを使って説明してきた。

Webサーバは、われわれからのtelnetを使ったリクエストに律儀に答えを返すところを観察した。そのときに、サーバは、無反省にだれにでも答えをかえしたことを思い出してほしい（ある意味戦慄的である！）。

Web で公開する内容（コンテンツという）は、特に何もしなければ、世界中のどこの誰でも手に入れることができるのである。この特徴は、大学でも、企業でも、「広報」とよばれる情報発信のタイプに非常にマッチするものである。一方、何らかの理由で会員限定のコンテンツをおかなければならない場合、誰がサーバにアクセスするかについて、気を配らなければならない。この理由には、組織の秘密に属するものもあるし、限定することで経済的価値を高める性質のものもある。

サーバが提供するサービスを、人や機械に対して選択的に提供することはよくあることである。この機構を一般に**アクセス制御**という。アクセス制御という概念は、このゼミナールの後半で扱う「セキュリティ」の主要なテーマのひとつである。

ところで、サーバが誰かからリクエストを受けたときに、どのような情報をそのリクエストから入手できるかは、アクセス制御の基本である。

Q. 前回だされた課題を思い出し、サーバがどのような情報をリクエストから入手できるかを想像せよ。また、どのような情報なら入手できないかを想像せよ。
Q. サーバが入手できる情報について、それがどこに記述されているか想像せよ。

以上は、とりあえず「想像」だけでかまわない。ただし、科学では、最終的に「想像」を「実証」するプロセスが求められる。

２．アクセス制御

この場合のアクセス制御とは、人間や機器に対して、ページにアクセス（読み書き）するかどうかを許したり、拒否したりすることになる。ここで「人間や機器に対して」という言葉に注意してほしい。実際には、サーバへは、「あるページに対するアクセス」の「リクエスト」が飛んでくる。サーバは、このリクエストから、それを出した「人間」や「機器」の情報を取り出さなければならない。

一般に「人間」や「機器」の情報を取り出し、本当であることを確認することを「認証」という。以下、認証がどのように行われるかをみてみよう。

２．１機器の認証

リクエストを出す機器はどうしたら認証できるだろうか？

Q.　「機器」として、思いつくものをあげてみよ。
Q.　前回の課題を思い出そう。目の前にある Mac の情報の何が kiku.itc.u-tokyo.ac.jp から返されてきただろうか。

大学が運用しているサーバの中には、「学内限定」と称しているものがいくつかある。これは、リクエストを出した機器の IP アドレスを認証して、それが学内の機器であることが確認できたら、それに対して応答しているのである。

Q.　IP アドレスがわかったとして、それが東大のものであるかどうかがなぜわかるのだろうか？
Q.　この認証に「ドメイン名」を使うにはどうしたら良いだろうか？　4 回目でやった DNS、また必要ならばコマンド dig (-x)を使って工夫してみよ。

Apache では、自分のページに対して、機器を認証してその結果でアクセス制御をかけることができる。

演習：

やりかたには 2 種類ある。httpd.conf で設定する方法と、フォルダごとに.htaccess(最初の”.”(ドット)をみのがさないこと)で設定するやり方である。ここでは後者のみを説明する。

1. 最初にフォルダを一つ作れ（名前は任意でよい）。
2. そのフォルダの中に以下の内容でファイル.htaccess を作成せよ。

```
order deny, allow
deny from (自分の端末の IP アドレス)
```

3. アクセスしてみよ。アクセスが拒否されたことを確認せよ。
4. 次の内容に置き換えよ。他人のサーバにアクセスを試みよ。

```
order deny, allow
deny from all
allow from (自分の端末の IP アドレス)
```

ここで重要なのは、アクセス制御のやりかたそのものではない。Apache には、機器の IP アドレスがわたってきて、それを認証する仕組みが提供されているということである。

ところで、前回の課題のときに、自分の端末の IP アドレスと、kiku から返された IP アドレスが異なっていることに気づいただろうか。

Q. 192.168 ではじまる IP アドレスを持つ機器には、一般にアクセスできない（ただし、例外はある）。その根拠を RFC の中から探し出せ（ヒント：special user IPv4）。

２．２　人の認証

機器の認証は、ネットワークの構成に密接に関係することがわかった。しかし、機器の認証だけでよいのだろうか？月曜日の 5 限だけ使える端末だということは、他の時間なら他の人もその端末を使う可能性があるということである。その端末を使っている人そのものの認証が重要になることはよくある。

計算機の世界では、この認証に伝統的に ID とパスワードによる認証を使ってきた（もはや説明は不要だろう）。これは、「あんた誰？」とサーバが聞くと「私は ID」と答え、それに対して「じゃ、証拠を見せてよ」と再度聞かれたときに自分しか知らないはずのパスワードを答えてその証拠とするやりかたである。ここで重要なのは、ID とパスワードのペアを作り出して記憶しておくことである。

演習：

１．http://www-sato.cc.u-tokyo.ac.jp/PKI-project/Komaba/Authにブラウザを使ってアクセスしてみよ。
２．同じ場所に telnet でアクセスしてみよ。
３．２つの違いから、ブラウザが返ってきた答えに対して何をしようとしたのかを想像せよ。

では、自分で ID とパスワードのペアを作り出してみよう。

１．ターミナルプログラムを立ち上げて、以下のコマンドを実行せよ。

```
% htpasswd –c .htpasswd (自分の名前)
```

２．先ほど（機器認証を要求するように）作成したフォルダの.htaccess を次のように書き換えよ。

```
AuthType Basic
AuthName “Web Seminar”
AuthUserFile /home01/uschuko/.htpasswd
                    (自分のホームの位置に合わせて適宜書き換えよ)
require valid-user
```

３．再度アクセスしてみよ。今度はどうなるか。

実は、Web サーバと、ブラウザの間には、以下のようなやりとりが行われている。

１．ブラウザは、普通に GET リクエストを出す。（確認せよ）

２．サーバは、Unauthorized であるエラーコードを返す。（確認せよ）

３．ブラウザは、ID とパスワードを用意して、再び Get リクエストを出す。

```
GET ページのパス HTTP/1.0
Authorization: BASIC bWl0ZW1vOmRhbWVkYXlvCg==
```

４．サーバは、（認証に成功したら）コンテンツを返す。たぶん、また拒否されるだろう。（これも確認せよ。）

上の Authorization: BASIC に続く呪文のような文字列が ID とパスワードをコロンをはさんで並べたものである。これを生成するには、以下のようにする。

```
% echo ID:password | nkf –MB
```

Q. BASE64 とよばれる方式について調べてみよ。

２．３　弱点

ここまでで、一段落したが、これで本当に大丈夫だろうか？たとえば、機器の認証にお

いて、自分が本来持っているアドレスと異なるアドレスがサーバに渡るのは珍しいことではない。これに対処するには、ドメイン名で機器を特定できるようにするとよいかもしれない。

また、ここまでの一連のやり取りは、すべて telnet を使ってできるものであったことに注意すべきである。telnet を使った通信はそのままネットワーク上を流れる。誰かが途中で自分とサーバのやりとりをのぞいているかもしれない（ただし、ちゃんとした組織ならば、誰でもこれができるわけではない）。のぞかれると、ID もパスワードも知られてしまい、誰か悪者が勝手にコンテンツをみるかもしれない。これに対処するには、通信の一部または全部を暗号化することが考えられるだろう。

３．おわりに

今回は、Web サーバのアクセス制御の話をした。アクセス制御に本質的な機器の認証と人間の認証の話もしたが、それぞれ対処の仕方は完全ではない。次回以降、現状での認証技術として安全性の面からも満足できる方法を探っていこう。SSL がキーワードになる。